Assessment of anomalies and extremes of surface temperatures and precipitation

This assessment originally used near-surface air temperatures measured by the Atmospheric Infrared Sounder (AIRS) on the Aqua satellite, and in particular the variable "SurfAirTemp", from the daily grid files.

Starting with January 2021, this assessment transitioned to using AIRS version 7 data. In the latter version, more conservative quality screening is applied during the retrieval of atmospheric profiles, causing reduced yield of valid profiles and consequently reduced yield of surface air temperatures, especially at high latitudes. To get more complete maps of extremes, a simple mean of surface air and skin temperatures, variables "SurfAirTemp" and "SurfSkinTemp", is analyzed.

The assessment is also using precipitation estimates from the Integrated Multi-satellitE Retrievals for GPM (IMERG), and specifically the IMERG Late Daily precipitation rates, variable "precipitationCal".

The assessment is performed once a month, when these data from the most recent month become entirely available. Thus, it is a monthly assessment using daily grid data. It is intended to be a brief report, steering readers attention toward some of the most notable patterns. In any given month, there are more patterns of anomalous persistent weather in the world than it would be practical to list here. Hence the bulleted list of the notable patterns is just a sample, and by no means is based on subjective preferences.

For brevity, let’s call the data from the most recent month the “current month”, and the data from the same month but from all past years of the base period the “reference month”. Statistical samples per spatial region are built separately for the current month, and for the reference month, and are used as input to the statistical analysis. Note, the reference month does not include data from the current month for a simple logical reason - we want to see if the local extremes in the current month have been observed in the past, i.e. the reference month, and if these extremes may be records. “Record extremes” should be understood strictly in the sense of the base period which for AIRS starts in September, 2002, and for IMERG in June, 2000.

For temperatures, the sample-defining spatial region is every original 1x1 deg grid cell from AIRS daily grid files. For precipitation, the spatial region is 0.5x0.5 deg area that contains 25 observations per day from the original 0.1x0.1 deg grid. Assuming 20-year observational record, this would define reference sample size of about 30 (days) x 19 (years) = 570 observations per spatial region for temperatures from AIRS. For precipitation, it would mean much more confident sample size of about 30 x 19 x 25 = 15000 observations per spatial region.

These numbers are good approximation of the maximum reference sample size that needs to be statistically analyzed for every spatial region where observations exist, to arrive at the final global maps of statistical results for all regions. Since every region forms an independent sample, in the case of near-surface air temperatures there are 360 (longitudes) x 180 (latitudes) = 64800 independent samples, and in the case of precipitation 720 x 360 = 259200 independent samples, to analyze.

Of course, natural processes affect the samples sizes, as well as the number of independent samples. For instance, precipitation is intermittent phenomena that is not present all the time within all 0.5x0.5 deg regions around the globe. Furthermore, IMERG algorithm is still being refined, and retrievals outside of the ±60 deg latitudes are sparse. Regarding the samples for temperatures, the main constraining factor is the cloud cover. Even though AIRS temperature retrieval still yields results for cloud fractions of up to roughly 90%, some areas with extensive and persistent cloud cover would yield significantly reduced sample size. However, unlike precipitation, the number of samples from high latitudes is mostly unaffected.

Considering the sample sizes, and the number of independent samples, relatively simple statistical approach is employed here. Every sample is sorted first. In the current month, for a given region, the highest and lowest 10% of the values are then easily extracted by subset index from that region sample, and the mean is taken separately for the highest- and the lowest-values subset. Taking the mean value, rather than the simple 10% cut-off, is just one of preferred choices because the mean adds more confidence in the extreme value, while at the same time smoothing unrealistic extremes. Thus found extremes are looked up in the reference sample for the same region.

Let's say the closest to the current highest 10% value is found at index i_{max} in the reference sample. Then, the frequency of occurrence of this and higher values in the reference past is $F_h = 1-i_{max}/N$, where N is the size of the reference sample. If $F_h<0.1$, that would mean that the currently observed extreme and higher values were less frequent in the past. For instance, if $F_h=0.01$ it means the current extreme and higher values were observed once in 100 reference observations, but in the current month were observed in 1 out of 10 observations. Hence, there is a 10-fold increase in the frequency of occurrence of the current extreme. If all values in the reference sample are below the current 10% extreme value, the latter is a current record for this region with respect to the refence period. The logic for extreme minima (coldest temperatures) is similar, after noting that the frequency of occurrence for this and lesser values is $F_l = i_{min}/N$. This analysis is done for every region where there are at least 10 valid values.

Having the values of F_h and F_l assessed for every region, it is possible to compute the percentage of the global area where $F_h<0.1$ (more frequent warm extremes in the current month), and $F_l<0.1$ (more frequent cold extremes in the current month). Note, these are outcomes from two different statistical assessments that refer to area, and should not be confused to necessarily add up to 100%. If the assessment outcome says 70% of the globe area reveal more frequent warm extremes in the current month, that only means that the rest 30% of the globe have either no change in the frequency, or saw more frequent warm extremes in the past reference (base) period. Another case where quick intuition is a bad advisor – the likelihood that same region will manifest increased frequency of occurrence of both warm and cold extremes in the current month is very real, and must not be dismissed either statistically or physically (IPCC, 2013).

In this assessment, an anomaly is the deviation of the mean of the current month from the mean estimate from the reference month. Anomalies are computed per region, with respect to the reference month (see above discussion of base periods, regions, and resulting sample sizes) for consistency with the assessment of extremes.

Symbolically, the anomaly Δy in every region can be expressed as:

$$\Delta y = \frac{1}{N_c}\sum_{i}^{N_c} c_i - \frac{1}{N_r}\sum_{i}^{N_r} r_i$$

Where:

c_i Valid observations in a region, of precipitation or temperature, forming the statistical sample for the region, in the current month.

N_c Number of days with valid observations in the region, in the current month. At best, Nc = 31.

r_i Valid observations in a region, of precipitation or temperature, forming the statistical sample for the region, in the reference month.

N_r Number of days with valid observations in the region, in the reference month. For 20-year reference sample, at best $N_r = 31*20$

The reference statistical sample *r* does not contain the current month's sample *c* .

In the case of AIRS near-surface temperatures, many regions may have small sample size, with large deviations from the mean. Since the full size for e.g. reference 20-year sample is at best about 600 observations per region (assuming 30 days in a month, all yielding valid samples), running Student t-test for confidence of the reference mean (Bendat and Piersol, 2010), for all regions (samples) is sufficiently fast. Hence, only those near-surface air anomalies are shown that reject the null hypothesis (no anomaly) with 95% confidence. Precipitation sample sizes per region can be substantially larger and at this point confidence test is not applied to them.

As a final remark, the definition of regions is underlined here on purpose to convey the intrinsic spatial averaging of satellite observations that in effect moderates the extremes and anomalies. By no means this assessment is in pursuit of sensational record-breaking extremes. For instance, in the case of AIRS near-surface air temperatures, in addition to their approximate attribution to the 2-m temperatures (Behrangi et al., 2017), the 1x1 deg region is a substantial natural smoother, and thus all temperature extremes reported here are muted. It is very likely local extremes on the surface, as experienced by people for instance, are much stronger in actuality.

References:

IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.

Julius S. Bendat, Allan G. Piersol, 2010: Random Data: Analysis and Measurement Procedures, Wiley, 4th Eddition, ISBN 978-0-470-24877-5

Ali Behrangi, Luke Chen, Van Dang, Eric J. Fetzer, Evan Fishbein, Fredrick W. Irion, Brian Kahn, Peter Kalmus, Evan Manning, Edward Olsen, Baijun Tian, R. Chris Wilson, and Sun Wong, 2017: AIRS V6 Test Report Supplement: Performance of AIRS+AMSU vs. AIRS-only retrievals, pp. 57-64